Markus Brunnermeier: So welcome back, everybody, and nice to see you again. Our webinar is organized by Princeton for everyone worldwide. We're very happy to have Luis Videgaray with us from MIT. Luis was the former finance minister and foreign minister from Mexico, and he is now an expert on artificial intelligence and how to regulate it. And we will talk about the race to regulate artificial intelligence today. Thanks, Luis, for being with us.

Luis Videgaray: Thank you, Markus.

Markus Brunnermeier: So let me give a few opening remarks, and then I'll pass on the phone or the microphone to Luis. Of course, innovation in the AI space is similar in many dimensions to other innovations, like technology and financial innovation, and medicine and pharmaceuticals. We have the FDA on aviation. And of course, there's a systemic risk component, like in. and finance. So the larger the AI will penetrate the economy, the more systemic it will become. Of course, on top of it, there's a singularity problem or singularity challenge, which might also lead to existential risk. And what's different about artificial intelligence is that it's self-perpetuating. So at some point, it might grow on its own without our intervention anymore. Another challenge is that we don't really understand how things are done. It's a black box, so the lack of explainability. And the question is, is this really different? Of course, when electricity was introduced, we didn't really understand fully electricity either. So that's the many other innovations where we don't understand fully initially and only later on what's going on. Of course, I guess it's a bigger challenge for artificial intelligence. So the different approaches to how to regulate artificial intelligence is a broad principles approach, or it can be very detail-oriented. And you also have to be, I guess, more flexible, because we don't know how innovation will play out. You can regulate at the product level, or you can regulate at the method level. And the different ways you can intervene, one is to slow down the innovation through a moratorium, say, over the next two years, we cannot make any innovation in artificial intelligence. And there was actually a letter sent out and signed by many scientists to slow down and not to have innovations for some number of years. There's a sandbox approach in regulation, where you say, I have some little space where we experiment with new innovations. And then only after we know better what the outcome is, we allow for it. FDA approval, you can think of this way as well. There's some explainability laws, where you say you have to be able to explain what you do. And there are certain directions you can push innovation through regulation. You can make it pro-worker, labor-augmented, like Asimov is pushing for. But you can also make it pro-safety or pro-freedom in many other dimensions. There's a connection between artificial intelligence. innovation and resilience, and that's some work I do with Daniel Chen. So resilience is about after how a system reacts after a shock, but it can also be a transition phase. It can be a shift instead of a shock. So you shift into a different regime. And the question is how quickly will this shift occur? What's the speed of innovation? And resilience is all about adaptability. How can society adapt to the new environment? And if the change is very fast, if the speed of innovation is very fast, the adaptability might not be fast enough to keep up with the innovation. So when the speed of adaptability is actually higher than the speed of innovation, then you might actually take advantage of these new issues or the new innovations

and you reach a superior steady state.

3:45
But when the speed of adaptability is lower than the speed of innovation, then there's a danger of adverse feedback loops and you might even hit the tipping point and the whole thing might be disintegrating. So the speed of adaptability relative to the speed of innovation in artificial intelligence is something to watch out for. And then, of course, the two policy responses. One is to slow down the innovation, or we can also speed up the adaptability, so make society more adaptable to changes and easily to speed up to that. Now, when you regulate AI, you might also regulate the whole industrial organization structure, the market structure. And there are two competing views. One is that typically we say we would like to have a competition policy. We want to promote competition in order to reduce rent extractions from a few leading players. So we would like to have many firms and free entry in order to reduce rent extraction. So from this point of view, often what we do is pro-competition policy, so competition policy, I.O. policy is pro-competition. On the other hand, if you emphasize instead of the rent extractions, much more the risk, existential risk from proliferation, you might want to have a few firms to control the proliferation of danger. So if we think about nuclear technology, we are happy that only a few nations can actually use it and this actually minimizes the proliferation of the technology because it might lead to more existential risk. So there's a tension there between on the one hand you want for rent extraction reasons or minimizing rent extraction, you would like to spread it as broadly as possible, free entry, or in order to contain proliferation of the dangers, you might like to have only a few companies who can really run this. And one way out of this might be that you have a layered approach. There's a first layer, the baseline, the base technology or a platform technology of, for example, large language models, that's where it's controlled by a few players. And then there's a second or third layer with more bespoke technology, which then is, you know, much more competitive. And of course, you still want to take into account that R&D requires some non-competitive elements, so there should be innovation in that space. But layer one would be more standardized between a smaller number of players for layer 1. But then it's also the question, where's the cutoff between layer 1 and layer 2 and layer 3? You know, how much should be done on the layer 1 space and how much should be done at the layer 2 and 3 spaces? Now the other question which often comes up, you know, that's all very critical, the large language models, they used a lot of the existing internet content in order to train the large language models, you know, starting from text but also movies and pictures and so forth. And the question is, should this be allowed or not, or should they pay for this language, for the content creators? And one argument is essentially from an economic perspective, you should actually create incentives to create content, so the initial providers of the content should be compensated for that. But if you think about it, the content was already created in the past, so all the costs are sunk, so essentially for incentives reasons, there is no reason really to compensate the providers of content for the past provision. It only should be for new content after, let's say, November 2022. That's from a pure incentive perspective, but from a fair rent-sharing perspective, you might want to say, okay, there should be some compensation for the creators of previous content, and that, you know, otherwise it leads, you know, to a lot of industry concentration. And, of course, for these large language models, you need billions of

dollars of initial investments to create this, but if you have to pay for the content, it might be even more expensive, so the entry might be even more difficult.

7:41
 So the answer depends, you know, how much industry concentration do you want? If they have to pay for the content, there will be fewer players able to enter, and it also depends how easily you can get financing. If you can get financing very easily, then actually it's more fair rent-sharing. If financing is very difficult to contain, you reduce innovation much more dramatically. Finally, there is, of course, an international race in AI regulation. There is a race to potentially create some national champions. That actually calls for less regulation. You might even want to subsidize artificial intelligence in a particular way. And you might also direct the artificial intelligence research more to military and other perspectives in light of geopolitics. And then the question is, what are the international bodies which are best to handle the international coordination of this regulation? Is it something like the Financial Stability Board, like the Bank of International Settlement has for the financial space? Or do we need some international ethics boards? To what extent do our values agree? Or can we have that? And that leads me to the poll questions. So the first question, the poll question, is AI regulation, should it be primarily focused on considerations of ethics, economics, AI safety or geopolitics? And the answers were 52% said ethics, only 10% based on economics. And AI safety is about 32%. And geopolitics only 6%. So the majority, even though most of them will be economists in the audience, said ethics is key in those considerations, followed by AI safety with 32%. And only economics and geopolitics play a very minor role. That's, to some extent, surprising given the audience. The second question was, should LLMs use existing content, internet content, can they use it for free or not? And 60% said yes, they should do it for free. And 40% they should say no, they should not be able to use it for free. So 60% said open AI, and they don't have to pay for existing internet content. The third question was, should we slow down AI innovation or not? And the answer for that was 28% say, yes, slow it down, but 70% no, don't slow it down. And finally, the final question is, who should, you know, the body, the international AI policy body, which leads to this coordination across the globe, and the G7 or OECD should do it because they have similar values, and that's about 16%. The G20 or United Nations, a broader range of a country should do it, 40%. Or should it be like more on the military side and the refugee side, Geneva Convention style, 30%, or should it be like the Financial Stability Board, the BIS style, that's 14%. So there might be other options as well, but the majority went actually to G20, United Nations, followed by a Geneva style, so 40% or 30% for the two leading guys. So with this, thanks a lot for your attention, and we're looking forward to Luis' perspective in the race to regulate AI. Thanks again, Luis.

Luis Videgaray: Thank you, Markus. Very honored to be here with you, and thank you to those joining. After your excellent introduction, I have very little to say, but I'll do my best to provide some additional ideas. I should say that, let me start with a conclusion. In the race to regulate AI, the only honest answer that I can provide in how to regulate AI is that nobody really knows. And I'll come back to that, because we are seeing so many different approaches trying different things. But it all comes back to the fundamental issue, which is that AI is very new, is rapidly

changing, and we don't really know how to create the guardrails, even though we're very much cognizant of the problems.

11:53

So I'm going to go through a few slides. I want to start by sharing a personal experience, something that happened to me back in 2017. Back then, I was the foreign minister of Mexico and we decided to promote an initiative in the United Nations around AI, to promote best practices, to share the knowledge, particularly focusing on emerging economies in less developed countries. And so here you have me, I was talking in the UN, but what is important here is this was my view. I was talking essentially to an empty audience and the response was not very enthusiastic. Actually, the reason why the initiative was rejected in 2017, remember, this is only seven years ago, this is not that long ago, was because the term artificial intelligence was considered not to be serious. It was considered to be talking about artificial intelligence. Artificial intelligence was, you know, a lot of people in the UN thought that it sounded more like the Terminator or movies. It's not really material for the UN. So things have changed and have changed a lot. Today, artificial intelligence policy is a major area of activity and of concern of not only policymakers, but society in general. So when we say artificial intelligence policy, what is it? Well, there is no consensus available, but we pretty much have a notion. We share people who discuss these issues. I think most would agree that it has something to do with a policy, public policy intended to modulate the impact of AI, including maximizing the benefits, which can be quite significant. So what are the tools of AI policy? There are many tools, but I would like to classify them into three broad categories. The first one is government spending. So governments can have significant influence and actually have had significant influence in the development of AI through its history by investing public money, taxpayers' money. It can be by promoting research and development. Let's not forget that most breakthroughs in AI until, say, about 12 years ago were happening in research institutions essentially with government money provided. Of course, education, the role of public spending is crucial, but also creating the infrastructure, not only connectivity, but power to enable the data centers is important. Incentives like subsidies and tax breaks. We see in America the CHIPS Act today subsidizing microprocessors for AI, for example. And of course, the governments can play a role as venture capitalists. This is something that doesn't happen a lot in the United States, but it's quite common in other jurisdictions, for example, in China. The second bucket is the use of AI in government. And some people forget how large governments are and how influential the adoption of particular tools of technology in government can be. So the use of AI to improve, hopefully, the delivery of public services, to aid decision-making in government, and even the military use of AI, can be very consequential in setting the path forward. So that's the second lever or second set of levers that the government can have. And of course, you have laws and regulations, coupled with international agreements and treaties. So for the rest of the conversation, I'm going to be talking more about laws and regulations, and a little bit about international agreements and treaties.

Markus Brunneremeier: Do you think that AI innovations are very different compared to other innovations, where the government is also providing some subsidies and orders? Are there

specific elements concerning AI?

Luis Videgaray: I think one of the key elements is that AI can be very, very transformative in government. AI can really change the workflow of how governments do many, many things, from tax collection, to permitting, to licenses, to benefit distribution. So the workings of government can change, and that's different from other technologies. But it's not the first time that has happened. I think the biggest challenge is, as you mentioned in your introduction, Markus, the deployment of AI is happening through a layered supply chain and the, what you called layer one, which are the base models or foundations, foundation models, that's becoming extremely capital intensive. And so you see, for example, universities have a hard time playing in that field. And it's only very large technology corporations and probably governments, the only players that can be there. So I think there's an interesting role here because of the very specific production function of AI, if we call it like that, where you have very high fixed costs and what appear to be so far economies of scale. So in particular that layer, and that calls probably from a different type of intervention than you would normally see in more conventional technology. So let's go now to the four camps of debate on AI policy. And these are the exact four camps that were in the first question in the poll. But let me just make some reflections out of this. First, this is just a summarization because we can slice and aggregate the different concerns. But the first thing that we should keep in mind is that AI presents many problems, not just one. And it's natural for some people to be concerned or some countries to be more concerned about particular problems. But the reality is that we have a multi-objective problem. AI policy is always dealing with more than one objective, in fact, many objectives. The other thing to keep in mind is that these objectives are not necessarily independent or orthogonal. There are some trade-offs involved. Even within the AI ethics camp, which is the older camp that started when I joined MIT back in 2019, a lot of AI policy was, the majority of AI policy was concerning what today we call AI ethics. And even within AI ethics, you can see some trade-offs and some tensions or dilemmas between objectives. For example, one of the most, one of the better known techniques for privacy protection, which is differential privacy, can exacerbate problems in biases and discrimination. And these are just technical trade-offs, but of course you have broader trade-offs like innovation versus consumer protection. And these camps, so the AI ethics camps, camp is now also in debate with the other camp that was popular in the poll, which is AI safety. AI ethics can be described, and again, this is a definition that can be controversial. It can be described as looking into the problems that we already know that are there. These are problems, not tomorrow's problems, but today are very real. The issue of bias and lack of explainability, persuasion for manipulation, those are problems that we have today. In AI safety, this camp is more concerned about the problems of the future, the problems of self-replication, the problem of autonomy, the problem of the challenges that come with safe improvement, and that could lead eventually even to existential risk. I should say that this is a camp that has gathered force much more recently. It's always been there, and we can see the literature that's always been there, but quite frankly, in academia, it was not very relevant until recently. Most of that, if you went back in 2020, to a room with computer scientists and spoke about existential risk, people would not be very interested in what we were saying then. Today

that has changed because the progress of AI as of late has been so impressive that now these concerns have become more relevant.

21:19
But there's a tension between the concerns of the present and the concerns of the future and some people, and I think with reason, claim in the ethics camp that AI safety can become a distraction and sometimes the distraction created actually to divert attention from AI ethics. So there's some tension there. Of course from an economist perspective, and I am an economist, there are many critical issues that are top of mind. Of course jobs is the most important one, but it's not the only one. Inequality, concentration as you mentioned Markus in your introduction, and even financial stability. For example, this is something that Gary Gensler now as you see a stock-debt length about, if you have one or just very few base models feeding off the application layers in the financial system, you might be creating monocultures and points of failure that are not detectable or not known. So there are many interesting questions, not only around jobs, but also in other concerns in economics. And of course, you have geopolitics. And there's clearly an awareness that AI will be very, very important in geopolitical presence and leadership. And it's become, in the eyes of many, a race. A few years back, maybe three years ago, four years ago, the consensus view was that there are two superpowers, the US and China, going head-to-head. I think more recently, it's clear that the US is more in the lead because of the development of generative AI and most of the development has happened so far in the US, not all of them, but some of the majority of them have happened in the US. And you see a lot of policy already trying to slow down the progress in competing countries, particularly in China, in the US-China relationship, we see things like the restrictions on the export of semiconductors, for example, that are, or the tooling for making the advanced semiconductors that actually are enforced now. And we now see much less communication between Chinese academia and US academia. But the geopolitics of AI is not only about US-China. You can think of it also as the divide between the makers of AI and the takers of AI. And most developing countries are essentially takers, particularly the foundation layer or the base layer. So there's a very important dynamic, and perhaps it's not the most, it's not necessarily the most talked about, but the reality is that most of the world is in a taker position, not a maker position around AI, and that creates tensions, of course. The first step towards creating AI policy were national strategies, started in 2017. Canada was the first country to have a national strategy. The US was only in 2019. By now, at least 50 countries have AI strategies, and the strategies are essentially non-binding, most likely high-level documents that set a path, set where we want to go, but are typically not very specific and definitely are not rules, are a good first step, but are only the first step towards defining policy. After strategies, the second wave of AI policy documentation became the AI principles documents.

Markus Brunnermeier: Luis, is there any impact? Can you see all these strategies had a significant impact in which direction AI went or was it just with little impact?

Luis Videgaray: I think the AI policy, the AI strategies have very little impact in terms of regulation, but they did have more impact in terms of government spending and in terms of the

use of AI in the public sector, the universities, public universities being created. I think some countries took the strategies more seriously than others.

25:30
Other countries are now revisiting their strategies and redoing the strategies because AI has changed so much. But I think in some countries, the strategies definitely had an impact, particularly on those two fronts, government spending and government adoption of AI.

Markus Brunnermeier: I have a question from Dennis. He was worried how much interest groups already play a role in developing these national strategies– do you see that, you know, certain... are cut off, or because certain interest groups don't want the development in certain directions, or do you think we're too early on this stage?

Markus Brunnermeier: That's definitely a serious issue, and I'll get to that in a moment when we talk about law. But because the national strategies are non-binding, and are directional governance, I don't think that interest groups were very much interested in that, even at the principal level, that I'll discuss in a second. But definitely in the law, and we'll talk, for example, about the European AI Act, you see interest groups very active. But I don't think that AI strategies were influenced by interest groups. We'll talk about laws in a minute. AI principles, essentially what they are, is agreeing on attributes. You can think instead of, if you, in your mind, substitute the word principles and say desires, these are clearly what the principles are. These are desired technical and socio-technical attributes that AI systems should have. And of course, it's a little bit harder to agree on principles than to agree on national strategies, each country individually, because these are typically international efforts, but there are already more than 100 documents of principles. Perhaps the most influential happened in 2019 with the OECD, and you can see a lot of it is about human-centered values, such as freedom, dignity, and autonomy, privacy, fairness, diversity, transparency, robustness, accountability. And these are principles that most people will agree on. The problem is that the documents that deal with principles, because they are non-binding, you can ask for having everything, but there is no trade-off reasoning involved. And that's the big difference between doing binding rules, as in laws, and doing principles, is that when you're doing actually binding rules, you need to deal with traders. So the laws around AI are just starting to emerge. Actually, the first country that enacted laws, is the first country I know about, was China. I think there's a misconception that China does not care about laws and the lack of regulation is an advantage. That's very far from true. China has quite sophisticated laws. Some of their laws pursue objectives that are more nuanced and quite detailed. But of course, Chinese laws also protect a particular form of government, which is different from democracy. So not necessarily the laws that are relevant in, the laws in China are relevant to other jurisdictions. The jurisdiction that has been perhaps the most comprehensive has been Europe. And Europe has been going through a very, I admire the goals and the process. The process has been quite comprehensive. It started in 20– the European AI Act, the first draft was published in April 2021, that is three years ago. And it's been a long debate, well-structured debate around what to do. The law will not come into effect until two years from now. So that's the challenge, but it has come with many challenges. The

first one is that this lengthy, thoughtful process has to be clearly surpassed by the speed at which the technology is evolving.

30:14
So the AI Act, the European AI Act was amended before being enacted. And there were many changes with the emergence of generative AI. But it's by far the most comprehensive approach to AI regulation. The United States-

Markus Brunnermeier: Do you think that's, so some people say Europe is moving so fast because it doesn't have so many AI companies. So it's easier to push a law through than in the United States. Or put it more cynically, so Europe has a lot of regulation but the US has a lot of firms doing AI.

Luis Videgaray: That's probably a cynical, too cynical of a view. I don't necessarily subscribe to that. First of all, Europe didn't move- fast. It took a long time and it's taking a long time. This is a law that will take a long time and I think it's based on a profound conviction around human rights and the role of government and the role of the state. That does not mean that the law is perfect and I actually have some significant concerns around the law and I'll talk about those in a minute, but I definitely respect the process and the goals. Whether the law will be effective or not, that's still an open question and I think there are reasons to be pessimistic about it, but again, it's an admirable process. The U.S. does not have such a process. There's activity in Congress, a lot of talk in Washington about AI over the last at least 18 months, but there are some proposals, many of them bipartisan, but it's unlikely that there'll be some federal regulation anytime soon. There are, however, pieces of regulation at the state level. For example, Colorado regulates the use of AI in insurance. The city of New York regulates the use of AI in hiring decisions, but still not necessarily a national framework. A lot of the AI regulation in the US will be defined by regulators. For example, the FDA, the SEC, the FTC, and also by courts. Judges and courts will play a very large role in the US in actually making decisions on how the law is applied to AI. Other countries, for example, in the developing world, you have the case of Brazil that introduced more than a year ago a very comprehensive law inspired by the European law. It has not been enacted, but you see action happening around the world. Let me just summarize a few and the previous slide was to give you just a flavor of how different the different jurisdictions are approaching these challenges. There are some key choices. The first one, do we need binding rules or not? Do we need AI laws or not? For example, the government of India last year came out and very explicitly said, we are not going to be regulating AI for the time being. The priority is innovation. So India will not be issuing any AI laws. The same happens, for example, in South Korea or in Israel. There's no intent to regulate AI where other jurisdictions are clearly doing the opposite, as we just saw. The second is, do we need one law or do we need many laws? Do we regulate horizontally in the sense that we have a crosscutting regulation like the AI Act, or do we acknowledge that the context matters and should we have a law for AI in healthcare, a law for AI in banking, a law for AI in education. This is an unanswered question. And I think that it's easier from a political perspective to grab attention towards horizontal laws, but perhaps vertical rules have a lot more practical merit than

horizontal rules. An interesting question is who's going to enforce the AI rules. Are you going to rely on the existing regulators or are you going to create new regulators?

Europe is creating new regulators, a new European AI office. In the US, it's very clear that it's gonna be, the burden is gonna be on existing regulators to enforce whatever rules are created or amended. And then the critical question to me is, are we regulating the inputs of AI or are we regulating the outcomes? And I'll give you an example. We all understand how critical data is. for AI. Essentially, AI today are just very sophisticated statistical tools that learn from data and then make predictions and actions based on that. So an approach would be to regulate the data and be less concerned about, for example, biases and fairness. If you have very good data that is representative, well-balanced and clean, regulating the data might be enough. That's, in essence, the approach that has been taken in Europe. There's an opposite approach, which is let's go for outcomes. Let's regulate what happens. It doesn't matter what kind of AI you're using, what kind of data, but if the outcome is, for example, in making a lending decision, a lending and writing decision, are you discriminating? Are you treating the customer unfairly? It doesn't matter if you're using AI or not, or what kind of AI, you cannot do it. So that's regulating the outcome that the consumer or Financial Protection Bureau in the US is taking the second approach. And it's clearly enforcing the pre-AI laws based on outcomes, regardless of what you're doing with that technology. I personally believe that, and I'll explain in a second, that regulating outcomes is probably more relevant than regulating the inputs, particularly because the technology is changing so much that it's very hard to put into law what exactly are you regulating if you focus on inputs.

Markus Brunnermeier: Is it also connected with the black box or the explainability issue that you don't know how the inputs translate into outcomes because it's a black box?

Luis Videgaray: That is true. I think that's part of the reason. And sometimes the outcome that you want to regulate is actually explainability. Going back to the example of a loan, in many countries, including in the US, the consumer that faces an adverse decision from a bank has the right to know why. if the bank is using an AI tool that is a black box and cannot explain why, and the only answer is, well, the algorithm is saying so, what the US authority has made clear to banks is that is not acceptable because we care about the outcome. We don't care necessarily about what you're doing behind it. So yes, definitely this is part of the issue. So let me just finalize with, I told you in the beginning, there is no clear answer on how to regulate AI, and countries are trying very different approaches, from no regulation to comprehensive horizontal regulation, to regulation in pieces, or to rely on the courts and the existing regulatory bodies. So why is it so hard to regulate AI? We've already discussed a few of the main ideas already, but the first one is very fast innovation. We can point to 2012 with the appearance of AlexNet as the moment where people realized that neural networks are for real and have a lot of potential. And since then, the acceleration of progress has been unprecedented. And even from November 2022, when ChatGPT was released and the world acknowledged that AI is very real, the progress since then has been substantial already. It's very hard to keep up. And the public

sector and the lawmaking process is not designed to be fast. And it should not be fast.

38:44
We don't want laws to be created in a hurry, but there's a very critical mismatch between the pace at which innovation is going, and it's accelerating, and the ability of the public sector to catch up. This is the first challenge. The second challenge we already mentioned, there are many objectives around AI policy and these objectives are not independent. There's no orthogonality in AI policy. The canonical, this is, I put on the screen the canonical technology policy example which is consumer protection through regulation versus innovation. And what we are seeing in the world is that countries are responding very differently to this challenge. But it's not supposed to just be laying out the principles or the desires, desired attributes of AI policy actually doing regulation and loss. You need to deal with the policy dilemmas and there are no straightforward answers to that. Then something that you mentioned already, Markus, in your introduction, AI today is built in a layered way. It's an AI supply chain. But what I want to bring to the attention is that this supply chain is absolutely global. So. If you are, for example, the government of Costa Rica and you're trying to regulate the use of AI in Costa Rica, you are dealing probably with a Costa Rica developer that is building on top, perhaps of an LLM or a base model that is in the US, and that is using data from Europe, that is probably using tools that were developed in Asia or in other countries, and this is happening in real time. So the jurisdictional authority of countries is challenged when you're dealing with a product that is not really happening under the authority of the country. Particularly for the emerging world, this is a challenge. I see a lot of countries trying to go the route of data sovereignty and compute sovereignty, but it's not clear that's an efficient way to do it when you have so pervasive economies of scale and network effects. I think that countries, particularly in the emerging world, are a little bit behind on catching up to this challenge, which has to do a lot with demanding guarantees, setting up clear liability rules from the companies that are deploying directly or indirectly their products into their jurisdictions. But it's very hard to regulate locally a product that is not only produced globally, but it's actually performing globally. The inference is happening beyond the borders of most countries.

Markus Brunnermeier: Can I ask you a related thought? Given that, do you think that smaller emerging market economies have a much bigger interest that there will be a global regulation, a global framework on that? And the big countries, what you call the AI makers, have less of an incentive to regulate the global scale and to take as the push for a global approach? Do you see this in the international policy world?

Luis Videgaray: Definitely, yes. But let me on something. For AI purposes, most countries are small. The majority of countries are small for AI purposes. So I think the majority of the world would clearly benefit from AI, from global AI regulation, but it's very difficult to get it done. And an example is lethal autonomous weapon systems. The Geneva Convention on Conventional Weapons started the talks in 2018, and the goal was to have something done by the end of 2019. It still hasn't happened. It's very difficult, even within a single domain, to get into an international agreement. And so far, there is really very limited success beyond principles on

having any actual rules being approved. But let me just go to the last two slides and just say something that we already mentioned.

42:58

Perhaps the most difficult thing about regulating AI is that it is a general purpose technology and to make the… I want to use an example to explain why this is so by looking at the different general purpose technology so we are very familiar with scissors. Scissors are a general purpose technology in the sense that we can use them for many things which are wonderful. We can use them in the kitchen, we can use them to manufacture things, clothing, of course surgeons use scissors in the operating rooms. Scissors can be wonderful but also scissors can be very dangerous. You can actually injure people with scissors, you can kill somebody with scissors, you can damage property with scissors but we don't have a law for scissors we don't have. you know, a claim, let's regulate scissors because they are very dangerous. We don't do that. What we do is we regulate liability. We regulate, we have laws, criminal laws against murder. We have liability rules in medical practice. So we regulate the, typically in general purpose technology, what you do is you regulate the outcome. It's very hard to imagine. It's very hard to imagine exactly what you are, what you're trying to do, to figure out what are the many future uses of a technology. And of course, the example of scissors is useful too, the analogy with scissors is useful on a certain point because AI is a much more complex technology than scissors and you have many different types of AI, many different types of models, but the core of the problem remains. It's a general purpose technology and trying to envision all the many uses and different ways. in which it can go wrong, it's very, very difficult. And finally, let me, you mentioned electricity, and I want to make this point based on the story of the battery. The man you see in the screen is Alessandro Volta, the Italian inventor of the battery, la pila. Alessandro Volta invented the battery in 1800. And it's very interesting that we only understood how batteries worked 60 years later when the emergence of the Maxwell equations. In fact, it happened that the explanations that Mr. Volta published about why his invention worked were completely wrong. And now we know that. It's not, it happens all the time. that human inventions are not properly understood in the beginning. And this is definitely the case with AI today. And this is the reason why many of the brightest minds in computer science, particularly in academic computer science, are devoting their time to try to understand the properties and the behaviors of these new, very large, very complex systems. And there are properties that we don't fully understand. We don't fully understand how, why these systems behave in a certain way, and why they do the things that they do. And it's very hard then to envision what the systems of the future, of the near future, of the next year, maybe by the end of the year, if GPT-5 comes before the end of the year, as expected, it's very hard to envision exactly how it will work. And not even the makers of the system know exactly why it works the way it works. Therefore, when you don't understand the workings of a product, it is very hard to regulate that product. And to try to regulate it from the inside, regulate the input approach, becomes very, very, very, very challenging. So this is the end of the talk. Markus, we'd like you to take a few questions.

Markus Brunnermeier: Yes, so thanks a lot, Luis. So one is the analogy with the scissors. But the scissors, I know what the scissors can do. I mean, you alluded to that. But AI, I don't even

know the possibilities of AI at this. It makes it even more challenging, I assume.

47:33
Luis Videgaray: Absolutely. The scissors point was only to make, using a simple example, to illustrate a general purpose technology. But AI is much more than that. AI, it's a new technology, it's rapidly changing. And we don't fully understand exactly why it works the way it does.

Markus Brunnermeier: The other question I have, you highlighted very nicely that different parts of the world react in terms of regulation very differently. But isn't this a good thing that different countries try different things and experiment in different forms of regulation, and then we figure out which one works the best and then we adopt the best? Isn't this what a rational approach would mean? It's actually good that we don't all do the same thing and make the same mistake. I 100% agree. I 100% agree with that. It also shows, and this should be humbling to humanity as a whole, that we don't know exactly how to do it. But the fact that we are trying different approaches in different places can bring a lot of learning. Hopefully, we have the ability to adapt once we see what works and what doesn't work, instead of doubling down on approach, because that can happen as well. But we're going to be able to learn a lot from the heterogeneity in the approach. And I want to ask you something. You know, in many countries, we're overburdened by bureaucracy. from the public sector and imposing on the private sector? And do you think artificial intelligence is a way to manage bureaucracy much more efficiently and reduce the burden of bureaucracy in a sense? Is this a way to resolve the bureaucracy problem? Because a lot of regulations can be handled more efficiently or handled by artificial intelligence in a sense. Actually, I'm very optimistic about the use of AI in government because of that issue. And I'm familiar with a case that I'm close to, which is the issue of government permitting for construction. This is an issue that every government in the world has. Construction permitting, nobody likes it, takes a long time. There's a lot of pieces to review. And the opportunity to have AI to automate and make that quicker and more transparent and less vulnerable to corruption is a huge win for it. Not only for the government, it's a huge win for the public sector, but for society in general. So I think that it's very challenging to actually change the workflows in government. It will take a lot of investment and some governments willing to be experimental. It's not gonna happen overnight, but it's very, very promising. And you can think the same around, for example, taxation or benefits distribution. There are so many areas in which we're just getting a driver's license. And there are so many areas where AI, the technology that we have today can be deployed into making those processes faster, more predictable. And I think we're gonna see that very soon. Again, it's not gonna happen immediately because it takes a lot of effort and investment to do that, but it's gonna happen.

Markus Brunnermeier: And some questions from the developing world or the audience. emerging countries. So Manuel Montes asks, where do you see the constraints in developing, you know, large language models and base layers of AI for the developing world? And a related question to that is, do the developing countries really need the developing AI's regulation on their own? So it's those particular things. For instance, what would a country like Lesotho benefit from expanding resources for AI regulation? Should they just let it play out and then adopt it

later on, the good things, and or should they be engaged in the debate?

Luis Videgaray: Both are brilliant questions. The base layer, there's a little bit of a debate of what kind of development will prevail. One school of thought is that this is going to be dominated by scaling, by going very, very big, where you have the next cohort of models that are probably going to cost over a billion dollars. So that excludes a lot of people around the world from doing that. And there are economies of scale and even network effects where you think that it doesn't make a lot of sense to have a lot of models. Of course, you have the opposite of that. You have the open source school and we're seeing recently some open source based models making significant progress. So that's a question that is to be debated but I think it will largely depend on which model prevails. Perhaps I'm inclined to think that closed source very large scale models probably have an edge but that's an open question. And that will largely drive how AI is adopted and what role do developing countries play in the AI space. I think the second question is also a very smart question. I don't think that all countries should regulate AI the same way. Particularly if you're small that is not going to be developing the AI or at least the base layer of AI, you're going to be more doing applications and customization of LLMs, for example, you want to be sure that you are regulating that process. And again, as I said before, if you want companies like Entropic or OpenAI are going to be providing their models for people in Lesotho to be doing a second layer or third layer of developments there, what are the guarantees that OpenAI and Entropic are providing to the developers and the people of Lesotho when their models are there? So it's a question about guarantees, it's a question about liability, and not necessarily a question about how are you going to localize data or localize computers, which I think a lot of the thinking in the emerging world is right now on those sovereign questions. But in a world dominated by economies of scale, it's not necessarily the right policy question.

Markus Brunnermeier: And do countries likely have to move fast in regulation, or can they wait a little bit and think carefully about it? What's the speed of regulation for them?

Luis Videgaray: I think that any country in the world should be thinking about this. The one thing that perhaps we can agree on is that ignoring the issue is not a good strategy. And I think that I wouldn't recommend, for example, the countries, small island countries in the Caribbean, to be doing their own AI laws. Probably not. But you need to be looking at how you connect to the AI supply chain, and what are the terms of service? How are you going to enforce those? And how are you going to protect consumers and developers in your own country? So I think that every country needs to have its own thinking and to address the problem, but acknowledging that context matters, and it's very different to regulate AI. U.S. that is to regulate it say in Saint Lucia in the Caribbean.

Markus Brunnermeier: Can I come back to closed source versus open source schools? One argument for closed source is that you know open source they don't correct for biases so discriminatory biases and all that and if you do open source essentially they might just discriminate. Do you see this concern as well or I mean of course there are a lot of arguments in

favor of open source but there's also and a closed source essentially they impose certain social norms on this on the large language models which you know they decide they decide you know the leaders of the like open eyes deciding what social norms are imposed on these models.

55:44
Yeah definitely the underlying model determines a lot of things but you can also particularly on things like biases and social norms there's a lot of adjustment and fine tuning that can be done in the second layer or third layer. For example, when you're doing retrieval-augmented generation, where you're constraining the type of answers and I don't know what one can do, you can do that locally. A lot of that can be done at the application layer, not necessarily at the underlying layer, but it's true. The fact that the big base models are localized, say, in the US, some of them are obviously in China, and now Europe, particularly France, is coming up with an interesting alternative. These models are trained mostly in English and not necessarily within the norms and don't have the same local cultures. Even local languages do not have the same degree of influence as the English language of the English culture or the Western culture are reflected there.

Markus Brunnermeier: Let me end with a very important question which came from an anonymous attendee. He said, what can we economists do, where can we add the most value to society in the AI space in addition to computer scientists? In other words, should all economists feel they have just begun to transition into machine learning, computer science? Or what do you see for economists the biggest added value?

Luis Videgaray: Well, I think there are many things to say about it. But obviously from a research perspective, AI should be a tool that everybody's using in economics. I think econometrics will never be the same. And the way that we use data and process data to learn about the economy should be very deeply influenced. From a policy perspective, I think that economists are not spending enough time looking into AI questions beyond the jobs question. I would love to see, for example, more industrial organization experts trying to understand from an economic perspective the production and function of AI. And what are the scenarios and what is the optical regulation based on what we know about the technology? We don't know everything about technology, but we know quite a bit of things. And computer scientists know quite a bit of it. And I don't see that happening as much. I think there's, and rightly so, there's a lot of attention to the jobs question and the skills and the displacement issue. But there are other, many other economics questions relating to AI in terms, for example, of stability or concentration or just industrial organization, vertical integration questions that I'm eager to see very smart economists around the world start to address.

Markus Brunnermeier: So we always end on a positive note. So if you look at, you know, the potentials of AI and the dangers, where are you on this spectrum? On average, I mean, on balance, the positive side given, or do you see the right regulations that the positives outweigh the negatives?

59:08

Luis Videgaray: I come from the developing world and I see the impact that AI can have in affecting the food supply, in affecting climate change for good, in affecting, you know, in improving healthcare outcomes, in improving education, and I think particularly in the developing world, the potential is huge. It's not there yet, but I'm very optimistic and very eager to see it. I acknowledge the dangers, I'm very conscious of the risks, but I think that sometimes we focus a little bit too much on the risks and overlook that incredible potential for improvement, particularly in the developing world.

Markus Brunnermeier: And you would say in the developing world, the advantages are even bigger than in the advanced economies.

Luis Videgaray: Of course, sometimes in the developed world, using AI or not, you already have a very good doctor, probably, and you already have a good technology, but if you are in sub-Saharan Africa, you probably will take a hard time to find a specialist and having some AI assistance clearly there has a lot more marginal value.

Markus Brunnermeier: Okay, let's stop at this positive note and I thank you a lot. Thanks a lot, Luis, for a great talk and we stay in touch and hopefully we can talk more about regulating AI and changing the world for the better. Thanks for doing this and talk to you soon. Thank you. Thank you.